Binding at LF vs. obligatory (counter-)coindexation at SS; a case study

1. Introduction *

T. Reinhart (1986) has recently challenged the "prevailing assumption in theoretical studies of anaphora that "the core issues in the case of pronominal anaphora (i.e. anaphora involving pronouns) are those of intended coreference with definite NP's" (p. 123) by arguing that "the core issue of the binding theory (or the syntax of anaphora) is that of bound-variable interpretation" *(ibid.)*. Independently of her work, D. Sportiche (1986), comparing the anaphoric/pronominal systems of English and Japanese, reached about the same conclusion that there are in fact *two* types of pronouns like *he/him/his* in English, the first one representing "pronouns as variables" and sharing with the strict anaphor *himself* the properties of Japanese *zibun,* and the second one corresponding to "referential pronouns". Both authors thus converged in treating alike strict anaphors and the first type of pronouns, henceforth archetypally *him-1,* and dealing separately with the "pragmatic" (T.R.) or "referential" (D.S.) *him-2.*

After presenting and illustrating part of their argumentation (§ 2), I will introduce some possible theoretical consequences and paradoxes due to a tentative introduction of R-expressions into the binding paradigm (§ 3). Next, I will show that the Reinhart-Sportiche basic distinction does carry over to the description of the so-called "reflexive" or "intensive" genitive *bere* of Northern Basque, thereby

weakening my former claim that Basque had no VP and was accordingly non-configurational: if *bere*-1 induces "sloppy-identity" interpretation (cf. Ross 1969) under clause coordination and deletion of recoverable material in the second clause, and is therefore a "bound variable" which must be dealt with at LF, *bere-2*, on the contrary, never induces such sloppy identity readings (for four of the five informants consulted) and can, moreover, be the specifier of the subject NP, in which case it must be coindexed with the object NP; consequently, either the s-structure is flat or non-configurational, and *bere-2* is also constrained by c-command, or the S-S *is* configurational, and the obligatory local coindexing of this pseudo-anaphor is *not* subject to c-command, contrary to an empirical claim made in passing by D. Sportiche *(op. cit.)* according to which "natural languages never seem to impose locality requirements not involving c-command" (§ 4). In § 5, I will study some dialectal and subdialectal variations concerning the necessarily coindexed *bere* and its pronominal counterpart *haren*, a series of facts which will lead me in the conclusion to also refute T. Reinhart's *(op cit.)* claim that "pragmatic" or "intended" coreference has nothing to do with "the syntax of anaphora" (§ 6): the opposition between *bere-1* and *bere-2* merely boils down to the fact that the latter must already be coindexed at s-structure, whilst the former will have to be operator-bound at LF.

2. Anaphors, Bound Pronouns, and Referential ones in English and Japanese

2.1. The notion "bound variable" as used here, after Reinhart (1986), covers both standard (sentential) anaphors, such as *himself* in (1a), and necessarily bound pronominals like *his* in (1b) or *her* in (1c): they behave alike, although the latter two are merely "free", in the technical sense, in their binding or governing category GC, whether this GC is defined as in Chomsky (1981) or Chomsky (1986), as is demonstrated by the necessary sloppy identity interpretation which obtains in the second clause in the three examples:

(1) a. Peter$_i$ saw himself$_i$ and so did John
         [i.e. John$_j$ saw himself$_j$ - not Peter$_i$]
     b. John$_i$ lost his$_i$ life, and Peter did too
         [i.e. John died, and Peter died too]
     c. Mary$_i$ lost her$_i$ way, and so did Jane
         [i.e. Mary got lost, and Jane got lost too]

On the other hand, when a pronominal is either not coindexed with any other NP, or is, but without being c-commanded by it, this pronominal is not only technically free: it functions like a "name" or a "referential expression", as in the following examples, where the sloppy identity reading is unavailable:

(2) a. John$_i$, likes him$_{*i,j}$, and so does Peter$_k$
         [i.e. and Peter$_k$ likes him$_j$ too, where i≠j, i≠k, j≠k]
     b. John$_i$ likes Paul$_j$, and so does Mary$_k$
         [i.e. and Mary$_k$ likes Paul$_j$ too]

(3)a.John$_i$'s mother saw him$_i$, and so did Peter
[i.e. and Peter saw John too]
b.John$_i$'s mother saw him$_j$, and so did Peter$_k$
[i.e. and Peter$_k$ saw him$_j$ too, whoever that person him$_j$ might be]
c.John's mother likes Paul, and so does Peter
[i.e. and Peter likes Paul too]

In the case of (3a), Reinhart speaks of "intended" or "pragmatic" coreference; but the whole paradigm (2)-(3) indicates that it would be more natural to use Sportiche's expression "referential pronouns" in cases like (2a), (3a) and (3b), because their reference is fixed once and for all in the first clause, rather than "variable" or depending on a local antecedent.

As is well known, there also are cases when both sloppy and strict or non-sloppy identity readings are possible; note that in the sentences (4a, b) below, the antecedent, a definite expression, technically binds the pronominal (outside its GC of course).

(4)a.John$_i$ asked Mary to shave him$_i$, and so did Peter$_j$
(i)[sloppy] ... and Peter$_j$ asked Mary to shave him$_j$ too
(ii)[non-sloppy] ... and Peter$_j$ asked Mary to shave him$_i$
[= John] too
b.Mary$_i$ likes her$_j$ neighbours, and Sue$_j$ does too
(i)[sloppy] ... and Sue$_j$ likes her$_j$ (own) neighbours too
(ii)[non-sloppy] ... and Sue$_j$ likes her$_i$ [= Mary's] neighbours too

(Naturally, the two distinct readings are only available when him, her above are referentially dependent on the subject NP: if we had him$_k$ or her$_k$ in the first clause, only the "referential" or strict identity interpretation would be possible).

A final case worth exemplifying is one in which the subject is a (universally) quantified expression; here, as is well-known too, if the subject binds the pronominal in the first clause, the non-sloppy reading is impossible in the elliptical one which follows: [1]

(5)a.Everybody$_i$ hates his$_i$ boss; therefore Paul does too
(i)[sloppy] therefore Paul$_j$ hates his$_j$ (own) boss
(ii)[non-sloppy] *therefore Paul hates everybody$_i$'s boss
b.Nobody$_i$ will ask you to help him$_i$; therefore Paul won't
(i)[sloppy] therefore Paul$_j$ won't ask you to help him$_j$
(ii)[non-sloppy] *therefore Paul won't ask you to help anybody

(1) The ungrammaticality of the non-sloppy interpretation can be reduced to a general principle forbidding variables to be left unbound at LF (a principle which derives in its turn from the Koopman & Sportiche Bijection Principle which will be discussed in 4.2.1.). Thus, (5a-ii) can be "translated" into:
(a) $[_{P\emptyset}[_{P1}$ for every x, x human, $[x$ hates x's boss]] $\Rightarrow [_{P2}$ Paul hates x's boss]]
where x('s) in P_2 has no operator to bind it. I am indebted to Ph. Nabonnand for this remark.

Leaving aside R-expressions for the time being, it follows that if the orthodox binding theory only distinguishes between anaphors, which must be *locally* bound (= by def. both coindexed with, and c-commanded by, another NP), and pronominals, which must be *locally* free (= by def. not bound) —but may be either non-locally bound, or (locally or extra-locally) coindexed with a non-c-commanding antecedent, or yet be "contextually", i.e. extra-sententially, coindexed— the foregoing discussion and data show that a tripartite distinction must be established within the anaphoric/pronominal system, as represented in (6) below, where the "bound variable" use of the pronominals is symbolized by *him-1*, and its (fixed) referential value is typified by *him-2*:

(6) Items   bound variable   submitted to a locality [+] /
                                antilocality [-] constraint
      himself   +   +   (anaphors)
      him-1   +   —   (bound
                                                             pronominals)
      him-2   —   —   (referential
                                                             pronominals)

We shall see in section 4.2.2 that the logically missing item, [-bound variable, + submitted to a locality constraint] does exist: it is precisely the Northern Basque *bere-2* alluded to in the introduction.

2.2. What renders Japanese particularly attractive here is that, according to Sportiche (1986), the lexical distinction established in this language is not one between anaphors like *himself* and pronominals like *him-1* and *him-2*, but a distinction between a bound variable item, *zibun*, which is not submitted to any (anti-) locality constraint, and a referential pronoun *kare* "he/him/his" (f. *kanozyo*), which *is*, on the contrary submitted to an antilocality condition. Sportiche's approach has obviously one great merit with respect to *zibun*: one is no longer forced to explain why or how either (a) *zibun*, is, in standard terms, either a strict anaphor which must consequently be bound in its GC, or a bound pronominal, i.e. an item which must be both free in its GC, and bound in a wider domain, or (b) it has no GC —as on Fiengo & Haruna's (1987) analysis.

Let us now illustrate the differences between *zibun* and *kare*, borrowing our first examples from this latter paper: [2]

(7) a. John-ga [[Mary-ga *zibun*-o semeta]   koto-]o   sitte-iru
         J.-NOM M.-NOM self-ACC blamed fact-ACC knows
         "John$_i$ knows that Mary$_j$ blamed self$_{i,j}$" [= F&H (2)]

[2] Although the aims and results of the two papers are obviously at variance with each other, and despite the fact that according to F&H *kare* is not a pronominal proper, but rather a deictic element, those differences do not affect the issues and problems I am trying to address here. I will not consider either the fact that *zibun* may also "be interpreted deictically as Speaker or Addressee according as the S[entence] is declarative or interrogative", as summarized by Keenan (1988: 133) —although this fact should ultimately prove that, in Sportiche's terms, *zibun* also has a "referential" value; see also F&H (*op. cit.*, ex. (8) and (9)) on this question.

    b.   John-ga [[Mary-no *zibun*-no e-ni   tuite-no
         NOM   GEN   GEN   picture-LOC about
         hyooka-ka]   karai koto-]o   sitte-iru
         evaluation-NOM severe fact-ACC knows
         "John$_i$ knows that Mary$_j$'s evaluation of self$_{i,j}$'s painting
         is severe" [= F&H (3)]

(8) a.   [[zibun-no zidoosya-ga kosyoo-sita]   John-]o
         GEN car-NOM   broke down (REL) J.-ACC
         watasi-no   zidoosya-ni noseta
         my [I-GEN] car-LOC   gave a ride
         "(I) gave a ride ro John$_i$ whose [lit.: self$_i$'s] car broke down in my
         [own] car" [= F&H (4)]
    b.   minna-ga   [zibun-ga/*kare-ga suki-na]   hon-o
         everyone-NOM z./k.-NOM   like-PART book-ACC
         katta
         bought
         "everyone bought the book which (s)he$_i$ [lit.: self$_i$] liked" [=
         F&H (6)]
    c.   John-wa [daremo-ga   zibun-o semeta to]   omotta
         TOP   everyone-NOM z.-ACC blamed COMP thought
         "John$_i$ thought that everyone$_j$ blamed self$_{i,j}$," [= Sportiche (6)]

The examples in (7) show that *zibun* is not necessarily locally bound (by a definite NP), contrary to *himself* for instance: it *is* locally bound when it takes *Mary* as its antecedent, but it is "long distance bound" or a "pronominal anaphor" when *John* is the antecedent.

The sentences in (8) show that *zibun* can also function like a (standard model) variable, because in (a) it is bound (probably via an empty operator) to the antecedent of the relative clause which contains it, and because in (b) it is bound by a quantified expression (compare (5)); (8b) also shows that *kare* cannot function like a variable, a fact to which I return. Finally, (8c) recapitulates *zibun*'s properties, since there it is either long-distance bound by a referential expression, or short distance bound by a quantified one.

That *kare* cannot, on the other hand, function like a variable —in either sense of the word— is corroborated by the next examples, borrowed from Saito & Hoji (1983: 247 (9)), in which, again, there is a contrast between the potential antecedents of *zibun* and *kare*.

(9) a.   John-ga [*zibun-ga/kare*-ga Mary-ni kirawarete-iru to]
         NOM self-NOM he-NOM M.-LOC is-disliked   COMP
         omoikonde-iru (koto)
         is-convinced   fact
         "John$_i$ is convinced that he$_i$ is disliked by Mary"

b.   daremo$_i$-ga   [*zibun*$_i$-ga/**kare$_i$-ga Mary-ni kirawarete-iru
       everyone-NOM self/he-NOM   M.-LOC is-disliked
       omoikonde-iru (koto) [cf. (8b)]
       COMP is-convinced
       "everyone$_i$ is convinced that he$_i$ is disliked by Mary"

Of course, *kare* in (9a) is grammatical; but it need not be bound by *John*, whereas *zibun* has to. In any case, *kare* is out in (9b) because its antecedent there is a quantified expression, not a referential one, just as in (8b): in both cases, *kare* would have been grammatical with an index different from the one of the subject NP.

Since there is no VP deletion in Japanese (Whitman 1987, § 3.2), the distinction between the bound variable or sloppy interpretation (i) of examples (4a, b) and the "pragmatic" or "referential" reading (ii) of the same examples must be illustrated in a distinct context. Saito & Hoji (1983: 257) provide us with such a context, starting with an analysis of (10) [their (32)] in English:

(10)   a.   Only John$_i$ thinks he$_i$ will win

(11)   a.   There is only one person x, x = John, such that x thinks
             that x will win.
       b.   There is only one person x, x = John, such that x thinks
             that John will win

As was also pointed out by Reinhart (*op. cit.*: 128), the truth values of these interpretations are different; to stick to S & H's example, we thus have, respectively, the following more transparent paraphrases of (11):

(12)   a.   Nobody$_i$ but John$_j$ thinks he$_i$/she$_i$ will win
       b.   Nobody$_i$ but John$_j$ thinks he$_j$/*he$_i$/*she$_i$ will win

Now, the interesting fact about the opposition between *zibun* and *kare* is that (13a) below *only* has the bound-variable interpretation (11a)/(12a), whereas (13b) only has ("aside from the irrelevant reading in which *kare* refers to someone other than John" [S & H (1983: 257)]) the fixed or referential interpretation (11b)/(12b):

(13)   a.   John-dake-ga [*zibun*-ga katu to]   omotte-iru
             only-NOM   z.-NOM win COMP thinks
       b.   John-dake-ga [*kare*-ga katu to] omotte-iru

Therefore, the traditional and restrictive definition of "variable" as referring to an expression necessarily bound by a quantified expression or an operator must be abandoned (provisionally), and replaced by Reinhart's wider one, given at the beginning of section 2.1, and according to which the properties of the antecedent are irrelevant, since it may also be a definite NP. From this point of view, *zibun* has

exactly the properties of a bound variable thus defined, and *kare*, those of a refe-rential pronoun —which can "pragmatically" (but need not) corefer to a definite NP only.

3. Interlude: Can R-Expressions be incorporated into the Reinhart-Sportiche framework?

3.1. From a purely descriptive point of view, we can summarize the results ob-tained up to now in the following table, where a correspondence between the En-glish and Japanese lexical entities examined above is established:

(14)	Items	English	Japanese
	anaphors	*himself*	*zibun-1*
	pronouns as variables	*him-1*	*zibun-2*
	referential pronouns	*him-2*	*kare*

Let us now consider a possible extension of (14), which would also include N-headed NP's or "R[eferential]-expressions", using the following binary features: B[ound] V[ariable] (in Reinhart's sense), and P[ronominal] in an intuitively Chomskyan sense:

(15)	Categories/items	English	Japanese	[BV]	[P]
	anaphors	*himself*	*zibun-1*	+	−
	pronouns as variables	*him-1*	*zibun-2*	+	+
	referential pronouns	*him-2*	*kare*	−	+
	R-expressions	*John*	*John*	−	−

Such a presentation raises at least two problems. The first one is theory-internal: it concerns the possible parallelism between lexical, i.e. phonetically realized, ca-tegories as classified in (15), and empty or unrealized categories; I will not address it here, first of all because there is no well-established theory of what Basque e.c.'s are or even look like and because, anyway, the facts to be discussed in sections 4 and 5 should concern every possible linguistic theory, and not only the GB approach.

The second problem has to do with the exact definitions of the binary features [BV] and [P] used in (15). As far as the positive value of the feature [BV] is con-cerned, there does not seem to be any difficulty: a [+BV] item is one which is necessarily bound, i.e. both coindexed with, and c-commanded by, some antece-dent NP. [+P] elements, on the other hand, either may be anti- (or extra-)locally bound, or must be locally free— in which case they may still be coindexed, but then, extralocally again, and/or without the NP they are coindexed with c-com-manding them. The obvious characterization then is: an item is [+P] if and only if it can be antilocally bound; if it happens to be, it will be either [+BV] at the same time, or [-BV]. These options are clearly illustrated by the two readings of,

say, example (4a): under the sloppy identity interpretation, we have *him-1* of (15), and, under the strict identity interpretation, we have *him-2*; note that in the latter case, the index of *him* in (4a) (or of *her* in (4b)) could have been *k*, (with $i \neq i, j \neq k$, and $i \neq k$ of course): as noted before, *him* (or *her*) would simply have been a referential pronoun, whose coreference with the (matrix) subject NP was a pure matter of chance or of "pragmatic" factors. The same results naturally obtain with the systematically different interpretations of *zibun(-2)* and *kare* in (13).

But there still remains one difficulty, linked to the [-P] value. We can of course derive the properties of strict anaphors like *himself* from the double fact that (a) they have to be bound, being marked [+ BV], and (b) they avoid being marked [+P] by not being able to be anti-locally bound. But what about the [-BV, -P] items? Note that the logical negation of [+BV] as defined above is *not* "must not be bound", but simply "need not be bound". (Besides, if [-BV] meant "must not be bound" *him-2* and *kare* would not even be allowed to exist at all). We therefore have a problem with R-expressions: being marked [-BV], they merely need not be bound, as we have just seen; moreover, being marked [-P], they still have the possibility of either not being bound at all (the correct result), or yet of being locally bound, a property which, I must insist, does *not* contradict their [-BV] characterization although it does contradict their essential property, which is stated by Principle C of the binding theory: R-expressions must be free (everywhere).

This naturally weakens the strength of (15) and of the associated definitions of the features [BV] and [P]: were it not for R-expressions, the distribution of anaphors and pronouns (as in English) or of (bound) variables and referential pronouns (as in Japanese) would have been entirely predictable, thereby rendering the Principles A and B of the B.T. superfluous.[3] Note however that if Principle C is perhaps too strong (as has often been argued in the literature), it is obvious that it will seldom be violated to the point of allowing R-expressions to behave like strict anaphors (see however Lasnik and Uriagereka (1988: 40) for a counter-example in Thai). Accordingly, we are confronted with a contradiction: the principles A and B seem to be (at least descriptively) necessary either as such, or modified, or yet as incorporated into the definitions of the [BV] and [P] features as used above, whereas Principle C should, according to many, either be dispensed with entirely within the syntax proper (as proposed by Reinhart 1986), or integrated into a "theory of discourse principle(s) for coreferential NP's" (as discussed and illustrated in Koster (1987: 353-4), where Reinhart's drastic distinction between syntax and

(3) This does not mean that the exact definition of (anti-)locality is straightforward; see for instance Koster (1985, 1987) and Manzini & Wexler (1987) for convincing arguments that locality does not only vary from language to language, but from lexical item to lexical item in the same language; cf. also van Riemsdijk (1985: 38):

> [..] while it is evidently an important advance to have a small set of binary parameters instead of virtually unlimited variation, it is obviously better to have no parameter at all. Whenever we observe differences in syntactic behavior among lexical items or classes of items, it is better to make the differences follow from properties of these lexical item[s] than to refer to these items and their properties in the principles of grammar.

This point of view is illustrated, as far as Basque is concerned, in Rebuschi (1987, 1988). See also Keenan's (1988: 134) remark: "As with Fijian, the Irish, Japanese and Turkish examples show that a theory of anaphora may not in general constrain the distribution of lexical items but only the range of interpretations available to these items".

pragmatics is blurred), or again submitted to parametric variation (as in Lasnik & Uriagereka, *op. cit.*); see also Milner (1986) for another attack against Principle C. In any case, the results obtained here go directly against all of this, since they render some independent version of other of Principle C more necessary than ever.

3.2. It seems to me that the heart of this contradiction lies in the very incorporation of R-expressions into (15): the [BV] and [P] features only concern "the syntax of anaphora *proper*" —to paraphrase Reinhart's words, i.e., more specifically, the anaphoric/pronominal systems of natural languages. In other words, something like a modified version of (15) *can* be maintained, provided the fourth line there is also occupied by an element which belongs to such an anaphoric / pronominal system —just as was suggested below (6), at the end of section 2.1.

Before I turn to showing that such an element does indeed exist in at least one natural language, let me make a final remark concerning the first three lines of (15): they exactly represent the three types of elements which Sportiche (1986: 370) entered in the "table 1" entitled "Locality conditions in English" —but which can also describe the two "uses" (from an English expert's viewpoint at least!) of Japanese *zibun* (see the first two lines of (14) and (15))— it being understood that (a) the "c-command requirement" and the "(anti-)locality conditions" represent an effort towards reducing a patent redundancy in (15), where the definitions of the [BV] and [P] features both included the word "bound", and (b) that the categories listed here are considered as coindexed with some antecedent, whether this is an obligatory constraint (1st column) or not (as assumed for the 2nd col.):

(16)		C-command required	C-command not required
	Locality condition	[strict] anaphors	*
	Antilocality condition	Pronouns as variables	Referential pronouns

Clearly, the question is whether it is possible to find an element filling in the gap or starred space in (16): if one is shown to exist, Sportiche's suggestion quoted in the introduction will have been falsified. Note that a "marked" relaxation on the c-command requirement for bound variables (such relaxations are discussed e.g. in Freidin (1986: 154-5) or Koster (1987: 326-7) cannot affect Sportiche's claim since what it says in essence is that there are (apparently) no languages possessing a lexical item which should be submitted to a locality (rather than an antilocality) condition and, at the same time, be (even optionally) coindexed without c-command being involved. It is this specific interpretation of Sportiche's statement which I will now challenge, on purely empirical grounds, by studying the properties of the so-called 3rd. p. "reflexive" genitive *bere* in the Navarro-Labourdin (henceforth NL) Basque dialect spoken in France (I will also use the expression "Northern

Basque", since the other northern dialect, Souletin Basque, does not seem to behave differently in this chapter of its grammar).

4. The two values of the Northern Basque Genitive *bere* "his/her"

4.1.1. The reason why I will concentrate here on this NL dialect (described in a traditional grammar framework by Lafitte (1962) —see especially pp. 93-3, §§ 208-211— and unfortunately ignored in Saltarelli (1988) is that in it *bere must* be coindexed with an argument of its own minimal clause (for conservative speakers) or of the matrix clause (for the remaining speakers, who probably are the majority today), contrary to the situation in the "Southern" or "Western" dialects spoken in Spain, where *bere* may be "discourse-bound". Now, in the latter case, there are two options: (a) *bere* may be either neutral as to the distinction between anaphors and pronominals, or else (b) it may be considered an anaphor bound by an empty Topic constituent (a solution which could be compared to Huang's (1984) proposal that referentially free zero pronouns are variables in certain languages). In either case, it is not clear how those dialects could help us to provide an item filling in the gap in (16).

4.1.2. Let us now turn to a few basic facts. Today's NL Basque (henceforth "Basque" only, unless otherwise specified) has two basic 3rd. p. "possessive" genitives: *bere* and *haren* (I take *beraren* to be an emphatic variant of *haren*; for more details, see Rebuschi 1988). On the face-value of such examples as the following:

(17) Jonek$_i$ ikusi du   *bere*$_{i,*j}$ / *haren*$_{*i,j}$ xakurra
      Jon-*k* seen AUX his   his   dog-Ø
      "Jon has seen his dog"

I have been led (e.g. in Rebuschi 1986, 1989) to consider *bere* an anaphor, and *haren* a pronominal, much as the specialists in various Indo-European languages distinguish between anaphoric possessives, such as Latin *suus*, Russian *svoj*, Danish or Norwegian *sin*, etc. and pronominal possessives/genitives like, respectively, *eius, jego, hans* and the like.

There is, however, one great difference between Basque *bere* and its Indo-European analogues, from now on noted °*sw*-. Thus, whereas °*sw*- is of course not only possible, but almost always obligatory in configurations like (18a) below, it is always impossible in cases like (18b),[4] *even when coreference is intended* (this being due, of course, to the fact that the object NP does not c-command the subject NP):

(18) a.   NP$_i$ [$_{VP}$ V [$_{NP}$ °*sw*-$_i$ N]$_j$]   b.   *[$_{NP}$ °*sw*-$_i$ N]$_j$ [$_{VP}$ V NP$_i$]

(4) Two facts must be noted. First, in Russian, *svoj* is not really compulsory —but only preferred— when the antecedent is 1st or 2nd p. rather than 3rd. Second, linear order and/or c-command are essential in these configurations; thus, there *are* violations of the filter forbidding Latin *suus* to appear within the subject NP of an independent clause, but this is only when the object NP has been fronted —and therefore c-commands the subject NP and *suus*, albeit from an A-bar position; see e.g. Riboni (1987) for many such literary examples.

On the contrary, the opposition between *bere* and *haren* is preserved in cases like (18b):

(19) a. [*bere*$_{i,*j}$ xakurrak] ikusi du   Jon$_i$
          his   dog-*k*   seen AUX Jon-Ø   .
          "his$_i$ dog has seen Jon$_i$"
      b. [*haren*$_{*i,j}$ xakurrak] ikusi du   Jon$_i$
          "his$_{j,*i}$ dog   has   seen Jon$_i$"

Recall furthermore that in the Northern dialects, *bere must* have an antecedent, so that it is out in (20):

(20) *haren*/*bere* xakurra hil   da
      h.   *b.* dog-ø   died AUX
      "his dog has died"

In my former work then, I derived from (20) that *bere* had to be bound, and from the grammaticality of (17) with *bere*, and that of (19a), that the subject and object NP's c-commanded each other —in other words, that Northern Basque was nonconfigurational, since the very existence of a VP node would have prevented the object NP from c-commanding, hence binding, *bere* in (19a), and from counter-binding *haren* in (19b) (the GC's for *bere* and *haren* being obviously identical, as demonstrated in Rebuschi 1987, 1988).

4.1.3. Of course, I fully realized that other anaphors, such as *nere/bere burua* "myself/himself", lit. "my/his (own) head" or *elkar* "each other" could *not* occur in subject position. Their behaviour, however, remains strange. For instance, whatever the person (1st, 2nd or 3rd) of the subject NP, *elkar* (or *nere/bere* etc. *burua*) is represented in the inflected verb form (generally an auxiliary) by the 3rd p. sg. prefix (*d-* in the present tense). Such facts can be illustrated by the following paradigm, in which only (a) is grammatical:

(21) a. guk^i elkar^j / gure burua^j ikusi dugu [d^j-u-gu^i]
          we-*k* e-Ø   our head-Ø seen we-have-it
          "we have seen each other/ourselves"
      b. *guk elkar/gure burua ikusi **gaitugu
                                  we-have-us
      c. *elkarrek^i/*gure buruak^i gu^j ikusi gaitu [gait^j-u-Ø^i]/**gaitugu

Another astonishing property of these anaphoric expressions is the following: in Basque, where *pro* may instantiate any of the three main grammatical functions or relations (subject, direct and indirect object), their person and number being, by the way, all represented in the inflected verb form, this "empty category" *seems* to be able to inherit from its discourse antecedent its [+anaphoric] property; here is

an example, borrowed from J. Hiriart-Urruty, a famous writer of the turn of the century (reprinted as J. H.-U. 1972: 85):

(22)   Nik$_i$ ez   dut   nahi agertu [nere$_i$ burua$_i$]; ez   dezaket ager
       I-*k* NEG AUX want show myself-Ø   NEG can't-it show
       "I do not want to show myself; (I) cannot"
       [lit.: "... I cannot [+ transitive] show *pro*"]

Therefore, there were grounds for believing that the equivalents of *myself, him-self* or *each other as anaphors* were to be dealt with at some level, or within some type, of representation which was *not* s-structure, and could well be K. Hale's (1982, 1983) l[exical]-structure, this structure being by hypothesis hierarchized (see also Mohanan 1984).[5] At s-structure, on the other hand, these NP's would behave like pronominals or R-expressions (whence the data in (21) and (22)). On the contrary, *haren* and *bere* would have to be taken care of at s-structure (note that being specifiers, they have no coarguments that could possibly bind them, contrary to *nere/bere burua* or *elkar*). Finally, this "dual" approach to binding in Basque had a nice independent consequence: if, in (22) for instance, [*nere burua*] as a whole is bound at l-s, and *nere* is bound at s-s, it follows that there is no *i/i filter violation in that sentence, since, presumably, this filter holds only of one type or level of representation at a time.

4.2. Let us now concentrate on *bere* (or *beren* in the modern language when the antecedent is plural). All the data utilized henceforth have been carefully check-ed with the five informants mentioned (in alphabetical order) in footnote (*). Except in a few cases, which will be duly pointed out when necessary, the judgments of four of them were remarkably convergent; I will call them the "standard" in-formants of NL basque, and will concentrate on their responses, leaving aside the radically different judgments of the fifth one until section 5.2.3.

4.2.1. First of all, as might be expected, *bere* functions like a bound variable whenever its antecedent is a quantified subject; this is demonstrated by the neces-sary sloppy-identity readings which obtain in the following sentences (the ending -*ak* in (23a) is not an ergative morpheme, but the irregular zero/absolutive marker of plural NP's):

(23)   a.   [emazte guziek]$_i$ [beren$_{i,*j}$ haurrak]   maite dituzte, bai eta
            woman all-*k*   their   children-Ø love AUX   yes and
            [gizon guziek ere/ eta gizon guziek ere bai
            man   all-*k*   too
            "[all women]$_i$ love their$_{i,*j}$ children, and all men (do) too"
            [i.e. and [all men]$_j$ love their$_{j,*i}$ children too"]

(5)   There is independent evidence that *elkar* and [*bere* (etc.) *burua*] must be dealt with at l-s or some other version of "argument structure": see Rebuschi (1987, 1988) for some Basque data, and Williams (1987) for a theoretical framework which I suspect might provide the means to deal with them.

    b.   [edozoin emaztek]$_i$ [*bere*$_{i,*j}$ haurra] maite du,   bai eta
        any   woman-*k b.*   child   love AUX yes and
        [edozoin gizonek] ere
        any   man-*k*
        "any woman$_i$ loves her$_i$ child, and so does any man"

But of course such examples do not entail that there is a VP at s-structure, since exactly the same effects obtain when the object precedes the subject:

(24) a.   [beren haurrak] [emazte guziek] maite dituzte, bai eta gizon
            guziek ere [same translation as (23a)]
      b.   [bere haurra] [edozoin emaztek] maite du, bai eta edozoin
            gizonek ere [same translation as (23b)]

On the contrary, the latter examples clearly show that *bere* as a bound variable mus be bound at LF, where QR [= Quantified-phrase Raising] has taken place. Note, incidentally, that Koopman & Sportiche's (1982) Bijection Principle is then apparently violated, since, after QR, the quantified subject "locally binds" (in K&S's terms) *two* variables: *bere* and its own trace. However, if (24a, b) are examples of Scrambling rather than Topicalization, the (hypothetical) d-structure linear order and hierarchy may be "reconstructed" at LF before QR applies, thereby preventing this violation of the Bijection Principle.[6] But then, if Basque *is* configurational at D-S and S-S, (24a, b) violate Saito & Hoji's (1983) requirement that a trace not be the antecedent of a pronoun or an anaphor which it does not c-command *at S-S*. All this appears more clearly in (25), where (a) the hypothetical configurational d-structure of (24a) —cf. (23a)—, (b) its derived s-structure, (c) its LF representation after QR but without, and then (d) with "reconstruction" or "descrambling", are respectively indicated (QP = quantified phrase):

(25) a.   [S [emazte guziek]$_i$ [$_{VP}$ [beren$_i$ haurrak] maite dituzte]]
      b.   [S [$_{NP}$ beren$_i$ haurrak]$_j$ [$_S$ [emazte guziek]$_i$ [$_{VP}$ t_j maite dituzte]]]
      c.   [S [$_{QP}$ emazte guziek]$_i$ [$_S$ [$_{NP}$ beren$_i$ haurrak]$_j$ [$_S$ t_i [$_{VP}$ t_j maite
         dituzte]]]]
      d.   [S [$_{QP}$ emazte guziek]$_i$ [$_S$ t_i [$_{VP}$ [beren$_i$ haurrak]$_j$ maite dituzte]]]

Therefore, either S&H's principle does not hold, or (NL) Basque is nonconfigurational; on the contrary, under the reconstruction hypothesis, K&S's Bijection Principle does hold good, but independently of the (non-)configurational nature of S-S, since it only applies at LF in the case under analysis; for some discussion whether LF might be "flat" too, see 5.2.3., where highly marginal facts are set forth.

---

(6) I take Scrambling to be a mere (Chomsky-)Adjunction to S/IP; therefore, it may either feed Reconstruction, if it is a syntactic phenomenon, or leave LF unaffected, if the movement takes place in PF (as J. Ortiz de Urbina has suggested —p.c.). In any case, the fact that Basque LF is configurational has never been challenged up to now, as far as I know (see however 5.2.3 and footnote 11 on this question).

4.2.2. That *bere(n)* can function *either* as a bound variable *or* a (necessarily —recall (20)) coindexed referential "pronoun" can now be illustrated by the ambiguity of the second, reduced clause in (26):

> (26) Jonek$_i$ *bere*$_i$ ama   ikusi du,   eta *Peiok*$_j$ ere bai
>   Jon-*k* b.   mother-Ø seen AUX and Peio-*k* too yes
>   "Jon$_i$ has seen his$_i$ mother, and Peio$_j$ (has seen his$_{i,j}$ mother) too"

Need we then distinguish between two *bere*'s? If the ambiguity of (26) invites us to do so, it seems to me that the following data leave no doubt whatsoever: when *bere* specifies the subject NP, its antecedent is the object NP, and, whatever their relative linear order, only the fixed or referential reading obtains in the reduced clause:

> (27) a. *bere* semeak Jon ikusi du,   eta Peio ere bai
>   b.   son-*k* J.-Ø seen AUX and P.-Ø too yes
>   "[his$_{i,*j}$ son]$_k$ has seen John$_i$, and (he$_k$/*[= Peio$_l$'s son]$_m$ has seen) Peio$_l$ too"
>   b. Jon, bere semeak ikusi du, eta Peio ere bai
>   ditto, lit. "Jon$_i$ his$_i$ son has seen e_i..."

Before drawing any theoretical conclusions from (27), note that in fact both sentences have two slightly distinct interpretations: one in which the NP left adjacent to the verb is not focused (or "foregrounded" in Uriagereka's (to appear) terms), and one in which it *is* focused; in the latter case, a marked variant is available in the Northern dialects in which the auxiliary immediately follows the focalized phrase, and precedes the lexical verb (see Lafitte 1962: 48, § 117/5 and Rebuschi 1983):

> (28) a. ? bere semeak, JON du   ikusi, ez   PEIO
>   b.   son-*k* J.-Ø AUX seen NEG P.-Ø
>   "his$_i$ son has seen JON$_i$, not PEIO"
>   b. Jon, BERE SEMEAK du ikusi, bai eta PEIO ere
>   "as for Jon$_i$, HIS$_i$ SON has seen him, and (as for) Peio too"

The question mark which precedes (28a) indicates that two "standard" informants in fact rejected the sentence —whilst the other two accepted it, but only with a fixed or referential reading as far as the second (elliptical) clause is concerned— this result being quite consonant with Horvath's (1986) remarks on comparable examples in English.[7] As for (28b), the four of them accepted it with a referential reading (or strict identity interpretation), and only one of them also *marginally* accepted it with a sloppy identity reading.

---

(7) What I mean here is that, whatever the exact nature of the S-S representation of sentences with a focused constituent (see Eguzkitza 1986, Ortiz de Urbina 1989 and Uriagereka (to appear) for very distinct proposals), (28a, b) is ungrammatical from a sentence-grammar viewpoint, although discourse considerations can render such structures licit —hence the fact that those sentences were judged acceptable by two of my four "standard" informants.

It then appears that, other things being equal (or dealt with by other modules of the grammar), when *bere* specifies the subject NP, its reference is typically fixed, even though it may happen to be technically bound by its antecedent: thus it is not only necessarily coindexed with, but also c-commanded by, the object NP *Jon* at s-structure in (27b) and (28b) —albeit from an A-bar position. However, c-command is clearly irrelevant here: when used as a referential entity, *bere* simply has to be locally coindexed at s-structure— if we allowed it to be coindexed only later on, at LF, we would transform it into a variable in the case of (28a), thereby allowing for a sloppy identity reading rejected by my informants; moreover, after the focused NP of (28b) has been raised and adjoined to S at LF, *bere* could no longer be bound by *Jon* there.

4.2.3. We may temporarily (see 5.2.2.) conclude that NL Basque has two distinct *bere*'s: *bere-1*, examined in 4.2.1, is a *bound variable:* it must be bound at LF, it is furthermore submitted to a locality requirement (for a slight qualification of this statement, see the discussion concerning the two subvarieties of NL Basque under 5.1.2): it therefore is a strict anaphor.[8] On the other hand, *bere-2*, illustrated

(8) Needless to say, such anaphors as *nere/bere burua* "my/himself" or *elkar* "each other", certain properties of which were described in section 4.1.3, always entail a sloppy-identity reading when the context is appropriate — compare (a) below to (1a) in the text:

    (a)   Peiok$_i$ [bere burua]$_i$ ikusi du   `[= d^i-u-Ø^i]$ eta Jonek ere bai
           P.-k   himself   seen AUX   and Jon-*k* too so
           "Peio has seen himself, and so has Jon" [not *and Jon has seen Peio too]
    (b)   Peio(k) eta Jonek elkar ikusi dute,   era Miren(ek)   eta   Arantxak   ere bai
           P.(-*k*)   and J.- *k* e.o.   seen AUX and M. (-*k*)   and Arantxa-*k* so   too
           "Peio and Jon have seen each other, and so have Miren and Arantxa"
           [*not*: and Miren and Arantxa have seen Jon and Peio]

Note however that *bere* (etc.) *burua* sometimes gets a literal interpretation; in such cases, this NP must be treated like a Referential expression, whence an ambiguity analogous to the one which obtains with (4) in English — or (26) in Basque; thus (a) has a counterpart (a'), in which only the indexing of the object NP has changed, apparently, but which has two interpretations rather than one:

    (a') Peiok$_i$ [bere$_i$ burua]$_j$ ikusi du eta Jonek ere bai
          "Peio$_i$ has seen his$_i$ head, and Jon$_j$ too (has seen his$_{i,j}$ head)"

It is this dual analysis of *bere burua* which may have led Hiriart-Urruty to build a sentence like (22), and induced Abaitua (1988: 199) to assert that the "reflexive expression" *b.b.* could function as the subject of its own clause; indeed, his example (30) (*ibid.*):

    (c) bere buruak izutu   zuen   Jon
       b.   head-*k* scared AUX[past] Jon

should *not* be interpreted as an instance of a true anaphor bound by the object NP, but as a case analogous to (27a,b), i.e. with *bere-2* and *bere burua* taken as an R-expression.

    Another remark is in order: various authors (Belleti & Rizzi 1988, Uriagereka 1989) have suggested that BT Principle A is in fact an "anywhere" principle, an assumption I am not following here. J. Ortiz de Urbina (p.c.) has however noted that in certain cases, the bound variable reading of *bere* cannot be obtained at LF, but at S-S, as in:

    (d) Jonek bere seme guz(t)iak maite ditu, eta Peiok ere bai
          "Jon$_i$ loves all his$_i$ sons, and so does Peio"

Here, the sloppy interpretation (which is not necessary, but only preferred, according to my standard informants) would not be available at LF, since the quantified object NP, which has undergone QR, has scope over the subject NP and asymmetrically c-commands it.

    Or does it? According to Reinhart's (1983: 23) precise definition of c-command (where A c-commands B iff there is a branching node C which dominates B and either immediately dominates A, or immediately dominates C', which in turn immediately dominates A, and C *and* C' *belong to the same category*), the subject

by (27) and (28), is always referential. Consequently, it can be described as a "pseu-do-anaphor", i.e. an element which is necessarily coindexed at s-structure —but either *not* submitted to c-command, if the s-structure is considered configuratio-nal— or yet *vacuously* submitted to c-command, is the s-structure is "flat". (Na-turally, the ambiguity of (26) can be explained in the same terms as those used to describe the two readings, sloppy and non-sloppy, of (4): we have *bere-1* in the former case, and *bere-2* in the latter).[9]

Accordingly, D. Sportiche's suggestion, already quoted in the introduction, that "natural languages never seem to impose locality requirements not involving c-command" is empirically falsified under a configurational analysis of NL Basque s-structure.

5. On Long Distance Binding and (Sub)Dialectal Variation

5.1.1. If (20) shows that *haren*, a real pronominal, need not be coindexed in the minimal or maximal clause which contains it, the examples (17) and (19b) show-ed that it fact it must be "counter-coindexed" locally exactly in the cases when *bere* must be locally coindexed. Moreover, in every example from (23) through (28), its referential index would have been not only distinct from that of the coargument of the NP it specifies, but fixed and stable, and disjoint, in the second, reduced clause too. (Under a configurational analysis of S-S, this fact of course strengthens my refutation of Sportiche's empirical claim, since *haren* is submitted to the same locality condition as *bere*).

But there *are* cases when *haren* may, however, be bound —extra-locally of course. Let us now consider such a case, (29) below.[10] What is noteworthy here is that both a sloppy identity reading, and a strict identity reading, are again possible, when *haren* is bound by the subject of the matrix clause (the third interpretation is only mentioned for memory's sake):

    (29)   Jonek erran du   [haren aita   jinen   dela],
             J.-*k*   said   AUX *h.*   father-Ø will-come AUX-COMP

NP would still c-command the object NP. It may thus be that owing to the distinct properties of *bere-1* and *bere-2*, NL Basque must force its anaphors to be bound at LF, whereas in those languages in which there are no items such as *bere-2*, there is no need for such a constraint.

(9) If the S-structure is flat, c-command will always apply vacuously to bound material at that level of representation. Hence the question: Why should *bere burua* "himself", *elkar* "each other" and *bere-1* also be bound at LF, when *bere-2* only has to be bound at S-S? (As far as I know, nobody has ever suggested that, within the same language, two lexical items would only differ in that one of them should be submitted to a given module in the syntax proper, and the other at LF —contrary to what has been suggested to account for certain cross-linguistic differences, as in Chomsky (1986a) for instance). Clearly, then, a configurational analysis of S-S should be preferred, but this in turn raises another question: Must we prefer substantive universals (in particular, the universal existence of a VP node) to such formal universals as c-command? See 6.2.2 for a highly tentative way of salvaging both.

(10) In the Southern *spoken* dialects, *haren* would normally be excluded in such a context, only *bere* (or the really emphatic pronominal *beraren*) being possible under coreference here; see Abaitua (1991) for a dis-cussion of this, and related, matters, and Saltarelli (1988: 96-133) for a presentation of *hura* (the absolute [-Ø ending] form of *haren*) vs. *bera* (whose genitive is *beraren*, just mentioned above, and *not bere*, which has no absolutive form, contrary to Saltarelli's assumptions) in the same southern dialects.

> eta Peiok ere bai
> and P.-*k* too yes
> [i]   "John$_i$ has said [his$_i$ father will come], and Peio$_j$ too (has said John's father will come)": bound *referential* value
> [ii] "Jon$_i$ has said [his$_i$ father will come], and Peio$_j$ too (has said his$_j$ (own) father will come)": bound *variable* reading
> [iii] "Jon$_i$ has said [his$_i$ father will come], and Peio$_k$ too (has said his$_j$ father will come)": fixed *referential* value

Therefore, there seem to be three types of *haren*'s, not only two (contrary to English *he/his/him*). *Haren-1*, which corresponds to the bound variable reading (29ii) above, also corresponds to English *he/his/him-1*. *Haren-2*, which is referential, cf. (29i) and (29iii), is the analogue of English *he/his/him-2*. But we must add the *haren* whose existence was recalled at the very beginning of this section; let us call it *haren-3*.

5.1.2. Before recapitulating our results, we must also take into account a subdialectal distinction which I have established elsewhere and is best illustrated by the fact that two of of my four standard informants also allow *bere* to be substituted for *haren* in (29); those who reject (30), i.e. more specifically *bere* in (30), simply seem to have a local domain for *bere* which is narrower than those who accept it there —a fact which is possibly linked to the parameter which consists in including, or not including, the notion of "accessible SUBJECT" in the definition of the GC's for anaphors, as suggested in Rebuschi (1988), following Yang (1983) and related work.

> (30)   ? Jonek erran du   [*bere* aita   jinen   dela]
>   J.-*k*   said AUX *b.*   father-Ø will-come AUX-COMP
>   "Jon$_i$ has said [his$_{i,*j}$ father will come]"

(Of course, as the indexes on *his* show, when judged acceptable, *bere* must co-refer to *Jon*.) But the interesting point is that the four informants, whether they accept or reject (30) as such, all admit on hearing (31) both a sloppy identity interpretation, and a fixed or referential one, for the second clause:

> (31)   ? Jonek erran du bere aita jinen dela [= (30)], eta Peiok ere bai
>   "Jon has said that his father will come, and Peio also has
>   [i]   (said that John's father will come)"
>   [ii] (said that his own [= Peio's] father will come)"
> just as they admitted two distinct interpretations for "bound" *haren* in (29i,ii).

5.2.1. Consider now the "restricted" or "conservative" subdialect (already attested in 16th century texts) in which (30) is out. *Bere* is then *strictly* local here (with the provisos mentioned in footnote 3), so that we can now modify (6) or (16) so as to account for the Basque data too. Indeed, a "basic" fact which was

overlooked, or simply left implicit, in both those tables is the level of representation considered. Now I think I have clearly established that *bere-1* must be bound at LF, as must of course all bound variables be. An obvious benefit is then that it is no longer necessary to have recourse to l-s (or argument structure) to deal with strict anaphors, as was suggested in section 4.1. (but see footnote 5 however). But there is more to it: (6) and (16) only allowed room for items wich were free to corefer or not extra-locally, so that the opposition between *bere-2* (the referential *bere*) and *haren-3* (which must be locally anti-coindexed) could not be taken into account.

I consequently propose the following table describing the distribution of *bere* and *haren* in the conservative variety of Northern Basque, leaving open the question of whether it would make sense in English and Japanese too (recall section 2) —but intuitively, I can think of no reason why it should not: to me, it is a sheer accidental fact that these languages should exhibit no lexical items submitted to a local (counter-)coindexation constraint at s-structure.

(32) *Binding and coindexation in conservative NL Basque*

	Binding at LF: obligatory	Coindexation at S-S:	
		obligatory	forbidden
Locality condition	*bere-1* (*himself, zibun-1*)	*bere-2* (*)	*haren-3* (*)
Antilocality condition	*haren-1* (*him-1, zibun-2*)	optional *haren-2* (*him-2, kare*)	

5.2.2. If we now turn to the "unrestricted" NL subdialect, the picture changes slightly, but not significantly (note that *bere-1'/2'* may be considered a fiction, if the exact definition of the GC of *bere* is parametrized, as suggested above):

(33) *Binding and coidexation in unrestricted NL Basque*

	Binding at LF: obligatory	Coindexation at S-S:	
		obligatory	forbidden
Locality condition	*bere-1*	*bere-2*	*haren-3*
Antilocality condition	*bere-1'/haren-1*	obligatory *bere-2'*	forbidden *haren-2*

Moreover, the variety of Southern Basque described by Abaitua (1991) could be reduced to a table in which the six spaces *could* be filled in by the "same" lexical

item *bere*, owing to the extremely extensive use of it in those dialects, but this remains to be carefully checked.

5.2.3. It is now high time I rapidly described the judgments of my fifth informant. Incredible though it may sound, he only accepts the sloppy identity interpretation *everywhere*, and in particular in such examples as (26), (27) and (28); however, he judged both (30) and (31) grammatical, admitting both strict and sloppy identity readings for the latter —just as for *haren* in (29). The only possible account I can find for such data is that:

(a) he does not possess *bere-2*, a fact which, if considered alone, would make his idiolect much more English/Japanese-like (hence unmarked?) than the standard informants', because, seen from a different point of view, *bere* is always an anaphor for him, or at least a bound variable (as mentioned in relation with (31), he possesses the *"bere-1"'* of (33));

(b) this idiolect is nevertheless highly marked in that it appears to exclude every possible hierarchical or configurational structure of S/IP *even at LF;* consider (27a) for instance: there, as has already been said, the object NP *Jon* is not necessarily focused, but may just happen to occupy the unmarked site for direct objects; however, only the sloppy identity reading obtains for the second clause: "his$_i$ son has seen Jon$_i$, and (his$_{j,*i}$ son has seen) Peio$_j$ too"; now this seems to imply that when the anaphor specifies the subject, it is the object which undergoes lambda-abstraction at LF:

(34)   [(Jon, (λx) [*x*'s son saw *x*]) and (Peio, (λy) [*y*'s son saw *y*])]

True enough, the deleted part of the second clause is, correctly, an alphabetical variant of its counterpart in the first clause. But we now have a double violation of the Bijection Principle, a phenomenon which, in its turn, demands to be accounted for. (Recall that for the other native speakers consulted, *bere* here merely had to be coindexed with a coargument of the subject NP which contains it, *at s-structure*, so that there is no need to have recourse to lambda-abstraction to account for this case of "pseudo-binding").

The solution that suggest itself is that LF too would be "flat" or VP-less for this fifth informant, so that the first clause of (27a) would have as its Logical Form representation:

(35)   (Jon, (λx) [*x* <object> (*x*'s son) <subject> has seen])

Obviously, more data should be gathered and analyzed in detail before such a highly marked hypothesis can hope to find significant independently justified support.[11]

_____

(11) I have mentioned elsewhere (Rebuschi 1986, 1989) that some "conservative" NL dialect speakers accept sentences like:

(a)   nor$_i$ maite du   bere$_i$ amak?
      who(m)   love   AUX   *b.*   mother-*k*
      "??/*who* does *his* mother love?"

6. Conclusions and New Problems

6.1. If the "syntax of anaphora" is to account for all the facts described in this paper (even leaving out those depicted just above, since they are due to one single informant) it is clear that its domain is much wider than the one the orthodox Binding Theory as such is able, and supposed, to cover. In particular, as many linguists (such as T. Reinhart and D. Sportiche) have shown long before me, this orthodox theory has nothing to say about the sloppy vs. strict identity interpretations of "bound pronouns" whose antecedents are definite expressions —a phenomenon which I have so much capitalized on here, and this should be enough to urge theoretical linguists to look for more appropriate tools.

Or should they? In fact, we already have them at hand: the only thing to do is to set apart *de jure* coindexation and c-command: "binding" as such is simply a misleading term, since the conjunction of coindexation and c-command only holds good at LF, where an anaphor (or anaphoric pronominal) must be technically "bound" by the (trace of a) variable. Accordingly, "binding" is not the tool we need to account for "intended" or would-be "pragmatic" coreference.

But must we, by way of consequence, concur with T. Reinhart in her claim that the latter has nothing to do with the "syntax of anaphora", and should rather be accounted for in terms of "pragmatic strategies"? I do not think so: the Basque data analyzed here have provided ample evidence that the opposition between *bere-2* and *haren-3 is* a syntactic matter, since both items *are* constrained by a locality condition. Therefore, when she writes (1986: 144) that when

> the grammar does not allow for bound anaphora, regardless of the placement of pronouns and antecedents [...] then, the hearer can infer nothing about the referential intentions of the speaker, and whether the NP's are intended as coreferential or not can be determined on the basis of discourse information alone,

she is only describing the situation which obtains in certain languages (perhaps the vast majority of them, but this is not to the point), which do not possess such lexical items as *bere-2* and *haren-3:* although the concrete use of these items has nothing to do with "bound variable anaphora", as we have seen, they do tell the hearer a lot about the speaker's intention to have his NP's corefer or not.[12]

---

which are usually predicted ungrammatical under either Koopman & Sportiche's Bijection Principle, or Saito & Hoji's constraint that traces may be antecedents of anaphors and pronouns only if they c-command them. But we have already seen that the latter proposal is hardly tenable in Basque. So, if we stick to the Bijection Principle, we should allow the same type of LF structure for (a) and for (35):

(b)   For which x, x a person, [x <object> (x's mother) <subject> loves]?

in which the object x A-binds the genitive x's contained within the subject NP.

But one difficulty remains: (a) above was judged grammatical by two of the four "standard" informants, who all reject the sloppy identity interpretation of (27a, b), and therefore have a grammar which excludes (35). I must confess I do not see how to renconcile these facts.

(12) L. Nash-Haran (p.c.) has kindly informed me that there is at least one more natural language which exhibits a pair of "anaphors" like *bere-1* and *bere-2:* Georgian, where the lexical item in question is *tavis.* Thus, parallel to (19a), we have:

(a)   tavis(-ma)   dzařlma   Vanos   ukbina
      self's (ERG)   dog-ERG   Vano-DAT   he-bit-to-him
      "his$_{i*j}$ dog bit Vano$_i$"

The main empirical contribution of Northern Basque therefore lies in its providing us with the opposition between *bere-1* and *bere-2,* an opposition which overtly demonstrates that the ambiguity of such sentences as (4a, b) is due to the fact that the lexical items *him* or *her* are either already coindexed (hence, given their reference) at s-structure, or not; in the former case, we get the strict identity reading in the second, elliptical, clause, but, in the latter case, the indexation is given later on, at LF, under lambda abstraction, therefore contributing the sloppy identity interpretation.[13]

6.2. From a theoretical point of view, at least two other important issues are raised by the present account of NL Basque properties. The first one has to do with Logical Form; the second one, which is concerned with the relationship between c-command and configurationality, will be taken up in 6.3.

---

although the word order in (b), where the antecedent preces *tavis,* is preferred:

> (b) Vanos$_i$ tavis(-ma)$_{i,*j}$ dzařlma ukbina
>   (id).

Contrary to Basque, however, the pronominal possessive *mis* "his" would be possible in both cases (compare (19b)); therefore, Georgian *mis* corressponds to *haren-1* and *haren-2,* but not to *haren-3* —a fact that has nothing disturbing about it, once it is acknowledged that locality conditions vary from one lexical item to another (see some references in footnote 3).

L. Nash-Haran also points out that Harris's (1981: 281, footnote 6) contention that "for many speakers, there is an additional constraint that *tavis* cannot occur in the subject" is a misled generalization due to the examples chosen:

> (c) *švils   bans   tavisi   deda   (d) *tavisi deda bans švils
>   child-DAT   she-bathes-himself's-NOM   mother-NOM   (id.)
>   "his$_i$ mother bathes the child$_i$"

In fact, the constraint would rather be that anaphoric possessives are generally avoided as specifiers of NP's which denote relatives or body-parts; moreover, the presence of the pronominal possessive *mis* in its canonical position, to the left of the head noun, fares just as badly; so, not only is *tavisi deda* out, but **misi deda* also is: the required construction is *deda misi,* a particular word-order which would not be acceptable in (a) or (b) with *mis(-ma)* substituted for *tavis (-ma).*

What is essential, in any case, is that when *tavis* is bound by a subject NP, then it is ambiguous (when the subject binder is a definite expression), leading to both sloppy and strict identity readings in the second clause, as in (e) below:

> (c) Ninom$_i$ tavis$_i$ amxanags   saati   čuka da Vanomac$_j$ igive   kna
>   Nino-ERG self's friend-DAT watch-ABS gave and Vano-ERG-too the-same-ABS did
>   "Nino$_i$ gave her$_{i,*j}$ friend a watch, and Vano did the same"
>   [i] ... and Vano$_j$ gave his$_j$ own friend a watch
>   [ii] ... and Vano gave him [= Nino's friend] a watch

whereas when *tavis* specifies the subject NP (an option which Harris allows for some speakers at least), then only the bound *referential* interpretation obtains, as in (f) [cf. (a) and (b)]:

> (f) Vanos$_i$ tavis (-ma)$_i$ dzařlma   magra ukbina
>   Vano-DAT   *t.* (-ERG) dog-ERG   badly   he-bit-to-him
>   da igive   ukna   Ninosac
>   and-the-same-NOM   he-did-to-her   Nino-DAT-too
>   "his$_i$ dog bit Vano$_i$ badly, and [his$_i$/*her$_j$ dog bit] Nino$_j$ [badly] too"

(13) As far as I have been able to make out, of all the alternatives to the general (albeit somewhat relaxed) GB framework still presupposed here, such as Higginbotham's (1983) "linking theory", Freidin's (1986) approach in terms of Theta-roles, or Chierchia's (1988) personal version of categorial grammar; none seems to be of any help in handling the *bere-2/haren-3* opposition; in particular all the theories which substitute F[unction]-command for C[onstituent]-command may not *in principle* account for the fact that *bere-2* is a pseudo-anaphor rather than a peculiar pronominal. As for Williams' theories, see 6.2.2 below.

6.2.1. The first problem related to my way of handling LF was alluded to at the end of footnote 8, and has to do with a possible parametricization of the binding systems of various languages; as I put it there, it may well be that it is just because NL Basque *bere* has two "uses" that its real anaphoric use must be dealt with at LF; but this is only a suggestion, which will have to await further justification.

6.2.2. Now here lies the most important issue: the very existence of LF as an autonomous level of representation has been challenged by several scholars lately (e.g. Williams (1986) and Koster (1987) between others). Thus, for Williams, LF would be a notational variant of S-S —with, however, a distinction to be drawn between S-S proper (where the scope of those items which have scopal properties corresponds to their actual site), and S-S', where the scopal index of those items has percolated upwards, a distinction that might be reinterpreted as S-S' being in its turn considered a notational variant of LF! But there is more to it: in Williams' system, predication intervenes as a mechanism whereby the VP which contains an anaphor is coindexed both with the subject and the anaphor, as in:

(36)   John$_i$ [saw himself$_i$]$_{VP;i}$

Consider now examples like the following:

(37)   a.   John$_i$ [read [his$_i$ book]] and Peter did too
       a'. John$_i$ [read [his$_i$ book]$_{(Np;i)}$]$_{VP;i}$ and Peter did too
       b.   Jonek$_i$ [[bere$_i$ liburua] irakurri du]   eta Peiok ere bai
            J.-*k* *b.*   book   read   AUX and P.-*k* too so
            [same meaning as (37a)]
       b'. Jonek$_i$ [[bere$_i$ liburua]$_{(NP;i)}$ irakurri du]$_{VP;i}$ eta Peiok ere bai

In (a) and (b), we have a representation of S-S proper, whereas (a') and (b') are possible instances of S-S', where a case of "vertical binding" is displayed; I assume that the object NP must also carry the index *i*, so that a "vertical chain" is constructed, without which no bound anaphora reading could be available for the tags (the elliptical second clauses). But now the following question must be asked: How is one to obtain the strict reading in the tags?

In the English case, it could be argued that the construction of the vertical chain is merely optional (after all, *his* is neither a typical anaphor —although it may sometimes necessarily be one, cf. (1b)— nor are *John* and *his* coarguments): if it takes place, we get the strict identity interpretation. But then, we have two distinct representations stemming out of the same one, viz. (38a), with a vertical chain, and (38b), without one:

(38)   a.   John$_i$ [read [his$_i$ book]$_{NP;i}$]$_{VP;i}$ b.   John$_i$ [read [his$_i$ book]]$_{VP;i}$

It follows that S-S and S-S' cannot be regarded as two innocent notational variants of each other.

Let us now turn to the Basque example. Recall that *bere* must of necessity be coindexed with a (term) NP, and the only available one is *Jon(ek)*; therefore, if one does not crucially distinguish between obligatory local coreference at S-S and binding proper at LF, one has to treat *bere* alike under both interpretations; above, in English, the strict identity reading could be thought of as resulting from some *accidental* coindexing of *John* and *his*, but this is no longer the case here: the coindexation is obligatory, local, and governed by c-command; consequently, bere *is* an anaphor, and nothing but a stipulation can prevent the vertical chain of coindexing from being built. There would thus be no way to devise a representation which makes the strict identity reading possible.

Looking at the problem from a slightly different point of view, we might say that in order to account for this non-sloppy interpretation, according to which Peio has read John's book, it is necessary to posit that *bere* has one of two distinct indices, say either i (the index of the subject *Jon*), or k (unspecified as to whether $i = k$ or $i \neq k$). The sloppy identity interpretation would obtain when *bere* has index i, because in this case it is indisputably an anaphor. But when it has index k, it behaves like a pronominal: no vertical binding chain may be built, so that the strict identity reading ensues. However, this mechanism requires yet another (type or level of) representation, where the identity of i and k is to be stated; let's call it S-S". Have we gained anything? I think not: if, assuming (as Williams does) that S-S and S-S' are non-distinct, we now have two different representations, S-S/S-S' on the one hand, and S-S" on the other (where the index k on *bere* is replaced by the subject's index i), instead of S-S and L.F. But since we have seen that there are good reasons to suspect that S-S and S-S' are anyhow to be dissociated, Williams' system (or my interpretation thereof) results in presenting us with three distinct representations, whereas the more classical system defended here has only two, given that in this system, there is no need to distinguish between S-S proper and S-S". In other words, if S-S' is taken to be a notational variant of LF, we are left with the results presented in 6.1: if the coindexation of *Jon* and *bere* takes place at S-S, we have a case of (bound) *referential* anaphora, i.e. an instance of *bere* as *bere-2*. On the other hand, if *bere* is not coindexed at S-S, it will have to be at LF, wherefrom we derive the sloppy identity reading. Parenthetically, note that on this approach, even though *binding* as such is not taken to be an "anywhere principle", at least *coindexation* must be considered to be one such principle.

6.2.3. The third problem connected with LF is that I argued in section 4.1.3 that the correct binding of certain Basque anaphors was probably best dealt with at the Argument Structure level (whatever its name may be). More work is therefore needed to either distribute binding over both LF and A-S, or to reduce the redundancies; again I must leave this topic for further research.

6.3. The final big theoretical problem raised here looks like a real paradox: it is the question alluded to in footnote 9, which concerns the apparent contradiction between recognizing a VP at S-S in Basque, and keeping c-command as one of the most fundamental structural relation in syntax, especially in local domains. Recall

(19a), or consider (39), with the object antecedent (binder?) in its canonical pre-
verbal position:

> (39) bere$_{i,*j}$ xakurrak Jon$_i$ ikusi du
>   *b.*   dog-*k*   J.-Ø seen AUX
>   "his$_{i,*j}$ dog has seen Jon$_i$"

The dilemma is this: either the obligatory coindexation of *bere* and *Jon* is gov-
erned by c-command, and there cannot possibly be any VP in (39), or this coin-
dexation is not governed by c-command, and the (unmarked?) hypothesis that Bas-
que has a VP can be maintained.

What is paradoxical about this alternative is precisely that the prototypical case
of c-command is the *asymmetric* c-command of an object NP by a subject NP, and
that it is rooted in an effort to rid structural dominance from precedence consi-
derations (Reinhart 1983).

Is there, then, a means of salvaging both? Recall that in the first half of the 80's,
there was strong disagreement about whether c-command was to be defined in
terms of the first branching node dominating the c-commander, or in terms of the
first maximal projection dominating it. In *Barriers,* however, Chomsky (1986b) ar-
gued that m[aximal]-command was the more general notion, the strict branching-
node c-command relation being restricted to binding; the argument ran as follows:
in (40) below, if m-command governs the relation between the trace *t* and the spe-
cifier, then a principle B or a principle C violation of the binding theory ensues;
on the contrary, in terms of strict c-command, there are no such problems:

> (40)   a. [$_{NP}$ its$_i$ [$_{N'}$ destruction t_i]]   b. [$_{NP}$ [the city]$_i$'s [$_{N'}$ destruction t_i]]

However, at about the same time, various scholars (among whom Fukui 1986
and Hellan 1986) developed the idea that "NP's" were D[et]P's really, conse-
quently, instead of (40), if we follow them, we have:

> (41)   a. [$_{DP}$ its$_i$ [$_{NP}$ destruction t_i]]   b. [$_{DP}$ [the city]$_i$'s [$_{NP}$ destruction t_i]]

in which case *t* neither c-commands, nor, more importantly, m-commands the DP
specifier. It therefore seems possible, today, to renounce strict (branching) c-com-
mand altogether, and to generalize m-command to binding relations too.

The relevance of the foregoing considerations is the following: given that there
are arguments in favour of the hypothesis that verbs govern their subjects in NL
Basque (Rebuschi 1989), it seems possible to tentatively posit that the Basque VP
is *not* a maximal projection (cf. Whitman 1987 for a similar proposal concerning
Japanese). Therefore, not only would the verb govern the subject NP, but the ob-
ject NP would also be allowed to m-command, hence to bind, the subject NP and
any material internal to it.

This would directly account for the properties of *bere-2* and for the fact that
the analogue *tavis(-2),* too rapidly described in footnote 12, belongs to the lexicon

of a language which also has many apparent non-configurational characteristics, namely, Georgian.[14]

Needless to say, the foregoing is highly tentative, and will require a lot of further research before it is (hopefully) ascertained. In particular, the peculiar properties of certain Basque anaphors (see 4.1.3.) might still require that the binding module apply to them at a level which is not S-S, because otherwise *elkar* "each other" or *gure burua* "ourselves" would improperly bind their antecedent in example (21a). A possible consequence of this approach would then be that the non-maximal character of VP's in Basque, Georgian, and Japanese-like languages is restricted to their S-structure —an independently welcome fact if binding at LF is not to suffer from the substitution of m-command for c-command.

Of course, yet other tracks might be pursued. In any case, the "strange" behaviour of *bere* in (19a) has now been reduced to less exotic considerations, and this, at least, is somewhat comforting.

7. References

Abaitua, J., 1988, *Complex Predicates in Basque. From Lexical Forms to Function Structures;* Ph D. thesis, Manchester U.

——, 1991, "Logophoricity and Emphatic Determiners in Basque"; in J. Lakarra & I. Ruiz (eds.), *Memoriae L. Mitxelena Magistri Sacrum* (San Sebastián, Anejos de *ASJU*, 14/2), 1013-1027.

Belletti, A. & Rizzi, L., 1988, "Psych-Verbs and θ-Theory", *NLLT* 6/3, 291-352.

Chierchia, G., 1988, "Aspects of a Categorial Theory of Binding"; in R. T. Oehrle, E. Bach & D. Wheeler (eds.), *Categorial Grammars and Natural Language Structures* (Dordrecht, Reidel), 125-151.

Chomsky, N., 1981, *Lectures on Government and Binding,* Dordrecht, Foris.

——, 1986a, *Knowledge of Language,* New York, Praeger.

——, 1986b, *Barriers,* Cambridge (Mass.) MIT Press.

Eguzkitza, A., 1987, *Topics on the Syntax of Basque and Romance;* Bloomington (Indiana), Indiana University Linguistics Club.

Fiengo, R. & Haruna, M., 1987, "Parameters in Binding Theory", in T. Imai & M. Saito (eds.), 107-128.

Freidin, R., 1986, "Fundamental Issues in the Theory of Binding", in B. Lust (ed.), 151-188.

Fukui, N., 1986, *A Theory of Category Projection and its Applications,* doctoral dissertation, MIT.

Hale, K., 1982, "Preliminary Remarks on Configurationality", in J. Pustejovsky & P. Sells (eds.), *NELS* 12, 86-96.

——, 1983, "Warlpiri and the Grammar of Nonconfigurational languages", *NLLT* 1/1, 5-47.

Harris, A. C., 1981, *Georgian Syntax. A Study in Relational Grammar,* Cambridge (U.K.), CUP.

---

(14) Incidentally, it would then not be a matter of sheer chance that both Basque *bere* and Georgian *tavis* should possibly be bound by a "term", i.e. the subject, or the direct or indirect (dative) object of their clause, but not by a non-term: other complements, and adjuncts, are PP's, i.e. maximal projections or categories which prevent the NP inside them from m-commanding anything outside of them.

Hellan, L., 1986, "The Headedness of NPs in Norwegian", in P. Muysken & H. van Riemsdijk (eds.), *Features and Projections* (Dordrecht, Foris), 89-122.

Higginbotham, J., 1983, "Logical Form, Binding, and Nominals", *LI* 14/3, 395-420.

Hiriart-Urruty, J., 1972, *Zezenak errepublikan* (edited by P. Lafitte), Oñate (Guipuzcoa), E.F.A.

Horvath, J., 1986, "Pronouns in Discourse and Sentence Grammar", *LI* 17/4, 759-766.

Huang, C. T. J., 1984, "On the Distribution and Reference of Empty Pronouns", *LI* 15/3, 531-574.

Imai, T. & Saito, M. (eds.), 1987, *Issues in Japanese Linguistics*, Dordrecht, Foris.

Keenan, E., 1988, "On Semantics and the Binding Theory", in J. A. Hawkins (ed.), *Explaining Language Universals* (Oxford, Blackwell), 105-144.

Koopman, H. & Sportiche, D., 1982, "Variables and the Bijection Principle", *The Linguistic Review* 2/2, 139-160.

Koster, J., 1985, "Reflexives in Dutch"; in J. Guéron, H.-G. Obenauer & J.-Y. Pollock (eds.), *Grammatical Representation* (Dordrecht, Foris), 141-167.

——, 1987, *Domains and Dynasties;* Dordrecht, Foris.

Lafitte, P., 1962, *Grammaire basque (navarro-labourdin littéraire)*, Bayonne, Editions des Amis du Musée basque & Ikas.

Lasnik, H. & Uriagereka, J., 1988, *A Course in GB Syntax. Lectures on Binding and Empty Categories*, Cambridge (Mass.), MIT Press.

Lust, B. (ed.), 1986, *Studies in the Acquisition of Anaphora I: Defining the Constraints*, Dordrecht, Reidel.

Manzini, R. & Wexler, K., 1987, "Parameters, Binding Theory, and Learnability", *LI* 18/3, 413-444.

Milner, J.-C., 1986, "Coréférence et coïndiciation: remarques à propos de l'axiome C", in M. Ronat & D. Couquaux (eds.), *La grammaire modulaire* (Paris, Minuit), 149-165.

Mohanan, K. P., 1984, "Lexical and Configurational Structures", *The Linguistic Review* 3, 113-139.

Ortiz de Ubina, J., 1989, *Parameters in the Grammar of Basque*, Dordrecht, Foris.

Rebuschi, G., 1983, "A note on focalization in Basque", *Journal of Basque Studies* 4/2, 29-42.

——, 1986, "Pour une représentation syntaxique duale: structure syntagmatique et structure lexicale en basque", *ASJU* 20/3, 683-704.

——, 1987, "Defining the three Binding Domains of Basque", paper read at the Conference on the Basque Language (2nd World Basque Congress), San Sebastián-Donostia, Aug. 31-Sept. 4; a slightly revised version is in *ASJU* 22/1, 233-241.

——, 1988, "A propos de quelques «universaux» de la théorie du liage", *Verbum* 11/2, 157-185; a modified version under the title "La problématique de la localité dans la théorie du liage" in P. Salaburu (ed.), *Sintaxi teoria eta euskara* (Vitoria: EHU/UPV), 121-141.

——, 1989, "Is there a VP in Basque?" in L. K. Marácz & P. Muysken (eds.), *Configurationality: the Typology of Asymmetries*, (Dordrecht, Foris), 85-116.

Reinhart, T., 1983, *Anaphora and Semantic Representation*, London, Croom Helm.

——, 1986, "Center and Periphery in the Grammar of Anaphora"; in B. Lust (ed.), 123-150.

Riboni, C., 1987, *La théorie du liage: sur l'anaphore en latin*, D.E.A. thesis, Univ. Nancy II.

Riemsdijk, H. Van, 1985, "Why Long Reciprocals don't Exist", *Theoretical Linguistic Research* 2/1, 37-45.

Ross, J., 1969, "Guess Who?", in R. Binnick et al. (eds.), *Papers from the Fifth Regional Meeting* (Chicago: Chicago Linguistic Society), 252-286.

Saito, M. & Hoji, H., 1983, "Weak Crossover and Move Alpha in Japanese", *NLLT* 1/2, 245-259.

Sertarelli, M., 1988, *Basque,* London, Croom Helm.

Sportiche, D., 1986, *"Zibun", LI* 17/3, 369-374.

Uriagereka, J., 1988, "La teoría del ligamiento", conferences at the 7th Summer Courses of the University of the Basque Country, San Sebastián, July 1988.

————, [to appear] "Variables in Basque and Governance", ms., MIT.

Whitman, J., 1987, "Configurationality Parameters", in T. Imai & M. Saito (eds.), 351-374.

Williams, E., 1986, "A Reassignment of the Functions of LF", *LI* 17/2, 265-299.

————, 1987, "Implicit Arguments, the Binding Theory, and Control", *NLLT* 5/2, 151-180.

Yang, D.-W., 1983, "The Extended Binding Theory of Anaphors", *Language Research* 19/2, 169-192.